Calcolo di Autovalori ed Autovettori: Il Metodo delle Potenze

Luca Gemignani
luca.gemignani@unipi.it

22 marzo 2018

Indice

Lezione 1: Generalità sul Metodo delle Potenze.

Il problema del calcolo di un autovalore e del corrispondente autovettore di una matrice $A \in \mathbb{C}^{n \times n}$ è affrontato mediante lo sviluppo di tecniche iterative che generano successioni $\{\lambda_k\}$ e $\{\boldsymbol{v}^{(k)}\}$ in modo da aversi

$$lim_{k \to +\infty} \lambda_k = \lambda, \quad lim_{k \to +\infty} \boldsymbol{v}^{(k)} = \boldsymbol{v},$$

con

$$A\boldsymbol{v} = \lambda \boldsymbol{v}, \quad \boldsymbol{v} \neq \boldsymbol{0}.$$

Tra queste tecniche si annovera *il metodo o i metodi delle potenze.* Nella forma più semplice essi generano a partire da un vettore iniziale $\boldsymbol{x}^{(0)}$, $\boldsymbol{x}^{(0)} \neq \boldsymbol{0}$, una sequenza di vettori $\{\boldsymbol{x}^{(k)}\}$ in accordo alla relazione

$$\begin{cases} \boldsymbol{z}^{(k)} = A\boldsymbol{x}^{(k)}; \\ \boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{z}^{(k)}}{\beta_k}, \quad k \geq 0, \end{cases} \tag{1}$$

con $\beta_k \in \mathbb{C}$ detto *fattore di scala o di normalizzazione.* Strategie usuali per la scelta del parametro sono $\beta_k = 1$, $k \geq 0$, che corrisponde al *metodo non normalizzato* e $|\beta_k| = \| \boldsymbol{z}^{(k)} \|$, $k \geq 0$, che corrisponde ai *metodi con normalizzazione.* Si osserva che se $\boldsymbol{z}^{(j)} \neq \boldsymbol{0}$ per $0 \leq j \leq k-1$, allora vale

$$\boldsymbol{x}^{(k)} = \frac{\boldsymbol{z}^{(k-1)}}{\beta_{k-1}} = \frac{A\boldsymbol{x}^{(k-1)}}{\beta_{k-1}} = \frac{A\boldsymbol{z}^{(k-2)}}{\beta_{k-2}\beta_{k-1}} = \frac{A^2\boldsymbol{x}^{(k-2)}}{\beta_{k-2}\beta_{k-1}},$$

da cui

$$\boldsymbol{x}^{(k)} = \frac{A^k \boldsymbol{x}^{(0)}}{\prod_{j=0}^{k-1} \beta_j}, \quad k \geq 0, \tag{2}$$

che motiva il nome attribuito al metodo.

Un'analisi delle proprietà di convergenza delle sequenze di vettori generate dal metodo (1) può essere condotta sotto le seguenti assunzioni.

1. La matrice $A \in \mathbb{C}^{n \times n}$ è assunta *diagonalizzabile*.

2. Gli autovalori di A possono essere ordinati in modo da aversi

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|.$$

   L'autovalore λ_1 è detto *autovalore dominante* di A.

3. Detti $\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(n)}$ gli autovettori corrispondenti agli autovalori $\lambda_1, \ldots, \lambda_n$ si ha che il vettore iniziale $\boldsymbol{x}^{(0)}$ rappresentato nella base di autovettori soddisfa

$$\boldsymbol{x}^{(0)} = \sum_{i=1}^{n} \alpha_i \boldsymbol{v}^{(i)} \text{ con } \alpha_1 \neq 0.$$

In particolare dalla relazione (2) per la proprietà (3) si ha

$$\boldsymbol{x}^{(k)} = \gamma_k^{-1} \sum_{i=}^{n} \alpha_i \lambda_i^k \boldsymbol{v}^{(i)}, \quad k \geq 0, \tag{3}$$

e

$$\boldsymbol{z}^{(k)} = \gamma_k^{-1} \sum_{i=}^{n} \alpha_i \lambda_i^{k+1} \boldsymbol{v}^{(i)}, \quad k \geq 0, \tag{4}$$

ove si è posto $\gamma_0 = 1$, $\gamma_k = \prod_{j=0}^{k-1} \beta_j$, $k \geq 1$. Ne discende il seguente.

Teorema 1.1. Se valgono (1), (2) e (3) allora le successioni dei vettori $\boldsymbol{x}^{(k)}$ e $\boldsymbol{z}^{(k)}$ generate dal metodo (1) sono ben definite ed inoltre si ha

$$\lim_{k \to +\infty} \frac{\gamma_k \boldsymbol{x}^{(k)}}{\lambda_1^k} = \alpha_1 \boldsymbol{v}^{(1)},$$

e

$$\lim_{k \to +\infty} \frac{\gamma_k \boldsymbol{z}^{(k)}}{\lambda_1^k} = \lambda_1 \alpha_1 \boldsymbol{v}^{(1)},$$

Dimostrazione. Da (3) e (4) segue che $\boldsymbol{x}^{(k)} \neq \boldsymbol{0}$ e $\boldsymbol{z}^{(k)} \neq \boldsymbol{0}$ per $k \geq 0$ essendo la componente di questi vettori lungo la direzione $\boldsymbol{v}^{(1)}$ non nulla. Per la relazione di limite si ha

$$\frac{\gamma_k \boldsymbol{x}^{(k)}}{\lambda_1^k} = \alpha_1 \boldsymbol{v}^{(1)} + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \boldsymbol{v}^{(i)},$$

e quindi da $|\lambda_i/\lambda_1| < 1$ si ottiene

$$\lim_{k\to+\infty} \frac{\gamma_k \boldsymbol{x}^{(k)}}{\lambda_1^k} = \alpha_1 \boldsymbol{v}^{(1)} + \boldsymbol{0} = \alpha_1 \boldsymbol{v}^{(1)}.$$

La relazione per $\boldsymbol{z}^{(k)}$ si prova analogamente. □

Il teorema implica il seguente risultato.

Teorema 1.2. Nelle assunzioni (1), (2) e (3) se $\boldsymbol{v}_j^{(1)} \neq 0$ per un certo indice j allora $\boldsymbol{x}_j^{(k)} \neq 0$ definitivamente, i.e., $\exists M : \boldsymbol{x}_j^{(k)} \neq 0 \ \forall k \geq M$.

Dimostrazione. Dal teorema precedente segue che

$$\lim_{k\to+\infty} \frac{\gamma_k \boldsymbol{x}_j^{(k)}}{\lambda_1^k} = \alpha_1 \boldsymbol{v}_j^{(1)}.$$

Posto dunque $0 < \epsilon < |\alpha_1 \boldsymbol{v}_j^{(1)}|$ si ha che $\exists M$ tale che

$$|\frac{\gamma_k \boldsymbol{x}_j^{(k)}}{\lambda_1^k} - \alpha_1 \boldsymbol{v}_j^{(1)}| \leq \epsilon, \quad \forall k \geq M,$$

e quindi

$$\boldsymbol{x}_j^{(k)} \neq 0 \quad \forall k \geq M.$$

□

Su questi risultati sono basate le tecniche di approssimazione dell'autovalore λ_1 e di un corrispondente autovettore.

Lezione 2: Approssimazione dell'Autovalore Dominante.

Un primo risultato per l'approssimazione dell'autovalore λ_1 segue immediatamente.

Teorema 2.1. Siano soddisfatte le assunzioni (1), (2) e (3) e sia $\boldsymbol{v}_j^{(1)} \neq 0$ per un certo indice j. Allora vale

$$\lim_{k\to+\infty} \frac{\boldsymbol{z}_j^{(k)}}{\boldsymbol{x}_j^{(k)}} = \lambda_1.$$

Dimostrazione. Dal Teorema (1.2) segue che la sequenza di approssimazioni è ben definita da un certo punto in poi. Per la relazione di limite si ha

$$\lim_{k\to+\infty} \frac{z_j^{(k)}}{x_j^{(k)}} = \lim_{k\to+\infty} \frac{\gamma_k \lambda_1^k \boldsymbol{z}_j^{(k)}}{\gamma_k \lambda_1^k x_j^{(k)}} = \lambda_1.$$

□

La determinazione dell'indice j può presentare alcune criticità computazionali. Nella seguente implementazione si sceglie j come l'indice della prima componente di modulo massimo del vettore $\boldsymbol{x}^{(k)}$, $k \geq 0$.

```
function [lam, it] = power1(A, tol, maxiter)
n=length(A);
x=rand(n,1)+sqrt(-1)*rand(n,1);
lam=+inf; err=+inf;
it=0;
while(err>tol && it<+maxiter)
    [s,i]=max(abs(x));
    xi=x(i);
    x=A*x;
    lam1=x(i)/xi;
    err=abs(lam-lam1);
    lam=lam1;
    x=x/s;
    it=it+1;
end
```

Per ovviare alla determinazione dell'indice un approccio alternativo considera la seguente funzione detta *quoziente di Rayleigh*

$$r(\boldsymbol{x}) = \frac{\boldsymbol{x}^H A \boldsymbol{x}}{\boldsymbol{x}^H \boldsymbol{x}}, \quad \forall \boldsymbol{x} \neq \boldsymbol{0},\ \boldsymbol{x} \in \mathbb{C}^n.$$

Vale infatti il seguente.

Teorema 2.2. Sotto le assunzioni (1), (2) e (3) si ha

$$\lim_{k \to +\infty} r(\boldsymbol{x}^{(k)}) = \lambda_1.$$

Dimostrazione. Posto $\boldsymbol{w}^{(k)} = \dfrac{\gamma_k \boldsymbol{x}^{(k)}}{\lambda_1^k}$ si ha

$$r(\boldsymbol{x}^{(k)}) == \frac{\boldsymbol{x^{(k)}}^H A \boldsymbol{x}^{(k)}}{\boldsymbol{x^{(k)}}^H \boldsymbol{x}^{(k)}} = \frac{\boldsymbol{w^{(k)}}^H A \boldsymbol{w}^{(k)}}{\boldsymbol{w}^{(k)H} \boldsymbol{w}^{(k)}} = r(\boldsymbol{w}^{(k)}),$$

e quindi per il teorema (1.2)

$$\lim_{k \to +\infty} r(\boldsymbol{x}^{(k)}) = \lim_{k \to +\infty} r(\boldsymbol{w}^{(k)}) = \lambda_1.$$

□

La seguente funzione implementa il metodo di approssimazione dell'autovalore dominante basato sul calcolo del quoziente di Rayleigh.

```
function [lam, it] = power2(A, tol, maxiter)
n=length(A);
x=rand(n,1)+sqrt(-1)*rand(n,1);
x=x/norm(x); z=A*x;
lam=+inf; err=+inf;
it=0;
while(err>tol && it<+maxiter)
    lam1=x'*z;
    err=abs(lam-lam1);
    lam=lam1;
    x=z/norm(z);
    z=A*x;
    it=it+1;
end
```

Il lettore confronti sperimentalmente i risultati restituiti dalle funzioni implementate per l'approssimazione dell'autovalore dominante delle matrici generate dal comando

```
A=gallery('tridiag', n, -1, 2, -1);
```

Il lettore dimostri che ler matrici così generate soddisfano le proprietà (1) e (2).

Lezione 3: Approssimazione dell'Autovettore.

Per l'approssimazione di un autovettore relativo all'autovalore dominante valgono risultati analoghi.

Teorema 3.1. Siano soddisfatte le assunzioni (1), (2) e (3) e sia $\boldsymbol{v}_j^{(1)} \neq 0$ per un certo indice j. Allora vale

$$\lim_{k \to +\infty} \frac{\boldsymbol{x}^{(k)}}{\boldsymbol{x}_j^{(k)}} = \frac{\boldsymbol{v}^{(1)}}{\boldsymbol{v}_j^{(1)}}.$$

Per l'approssimazione mediante il risultato dipende dalla stategia di scelta del fattore di normalizzazione.

Teorema 3.2. Nelle assunzioni (1), (2) e (3) con $\beta_k = \parallel \boldsymbol{z}^{(k)} \parallel$ vale

$$\lim_{k \to +\infty} \parallel \boldsymbol{z}^{(k)} - r(\boldsymbol{x}^{(k)})\boldsymbol{x}^{(k)} \parallel = \lim_{k \to +\infty} \parallel A\boldsymbol{x}^{(k)} - r(\boldsymbol{x}^{(k)})\boldsymbol{x}^{(k)} \parallel = 0,$$

Dimostrazione. Poichè $\parallel x^{(k)} \parallel = 1$ si ha che

$$\lim_{k \to +\infty} \parallel A\boldsymbol{x}^{(k)} - r(\boldsymbol{x}^{(k)})\boldsymbol{x}^{(k)} \parallel = \lim_{k \to +\infty} \frac{\parallel A\boldsymbol{x}^{(k)} - r(\boldsymbol{x}^{(k)})\boldsymbol{x}^{(k)} \parallel}{\parallel x^{(k)} \parallel}.$$

D'altra parte posto comde sopra $\boldsymbol{w}^{(k)} = \dfrac{\gamma_k \boldsymbol{x}^{(k)}}{\lambda_1^k}$ si ha

$$\lim_{k\to+\infty} \frac{\| A\boldsymbol{x}^{(k)} - r(\boldsymbol{x}^{(k)})\boldsymbol{x}^{(k)} \|}{\| x^{(k)} \|} = \lim_{k\to+\infty} \frac{\| A\boldsymbol{w}^{(k)} - r(\boldsymbol{w}^{(k)})\boldsymbol{w}^{(k)} \|}{\| w^{(k)} \|} = 0$$

per la continuità della norma. $\square$

Dal teorema segue che nei metodi delle potenze con normalizzazione il vettore $\boldsymbol{x}^{(k)}$, $k \geq 0$, fornisce un'approssimazione dell'autovettore relativo all'autovalore dominante. La seguente modifica della funzione `power2` incorpora il calcolo dell'autovettore e restituisce in output una misura del residuo corrispondente.

```
function [lam, x, res,  it] = power2_mod(A, tol, maxiter)
n=length(A);
x=rand(n,1)+sqrt(-1)*rand(n,1);
x=x/norm(x); z=A*x;
lam=+inf; err=+inf;
it=0;
while(err>tol && it<+maxiter)
    lam1=x'*z;
    err=abs(lam-lam1);
    lam=lam1;
    x=z/norm(z);
    z=A*x;
    res=norm(z-lam1*x);
    it=it+1;
end
```

Il lettore confronti sperimentalmente sugli esempi introdotti precedentemente l'errore di approssimazione dell'autovalore e dell'autovedttore corrispondente.

Lezione 4: Varianti del Metodo delle Potenze.

Il metodo delle potenze può essere modificato per permettere l'approssimazione dell'autovalore di modulo minino. Detti $\lambda_1, \ldots, \lambda_n$ gli autovalori di A ed assunto che

$$|\lambda_1| \geq \ldots \geq |\lambda_{n-1}| > |\lambda_n| > 0,$$

allora per gli autovalori $\eta_i = 1/\lambda_i$ della matrice inversa A^{-1} vale

$$|\eta_n| > |\eta_{n-1}| \geq \ldots \geq |\eta_1|.$$

Ne segue che il metodo detto *delle potenze inverse*

$$\begin{cases} A\boldsymbol{z}^{(k)} = \boldsymbol{x}^{(k)}; \\ \boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{z}^{(k)}}{\beta_k}, \quad k \geq 0, \end{cases} \tag{5}$$

permette nelle assunzioni usuali di determinare approssimazioni di λ_n e di un corrispondente autovettore. Analogamente se è disponibile una approssimazione iniziale η di un autovalore $\lambda = \lambda_i$ di A tale per cui

$$|\lambda - \eta| = |\lambda_i - \eta| < \min_{j \neq i} |\lambda_j - \eta|,$$

allora il il metodo detto *delle potenze inverse con shift*

$$\begin{cases} (A - \eta I_n)\boldsymbol{z}^{(k)} = \boldsymbol{x}^{(k)}; \\ \boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{z}^{(k)}}{\beta_k}, \quad k \geq 0, \end{cases} \tag{6}$$

permette nelle assunzioni usuali di raffinare l'approssimazione di λ_i ed eventualmente di un corrispondente autovettore.

Sotto opportune ipotesi il metodo delle potenze può essere esteso al calcolo di tutti gli autovalori dello spettro di A. Per fissare le idee assumiamo che A sia una matrice simmetrica. È ben noto che in questo caso A è diagonalizzabile ed inoltre autovettori relativi ad autovalori differenti sono ortogonali. Supponiamo inoltre che per gli autovalori $\lambda_1, \ldots, \lambda_n$ di A valga

$$|\lambda_1| > |\lambda_2| > |\lambda_3| \geq \ldots \geq |\lambda_n|.$$

Il lettore verifichi che la matrice $B = A - \lambda_1 \boldsymbol{v}^{(1)} {\boldsymbol{v}^{(1)}}^T$ ha autovalori $\lambda_2, \ldots, \lambda_n, 0$ e pertanto autovalore dominante λ_2. Ciò suggerisce la possibilità di procedere con l'approssimazione di λ_2 e di un corrispondente autovettore applicando il metodo delle potenze alla matrice

$$B = A - \tilde{\lambda}_1 \tilde{\boldsymbol{v}}^{(\mathbf{1})} {\tilde{\boldsymbol{v}}^{(\mathbf{1})}}^T,$$

dove $\tilde{\lambda}_1$ e $\tilde{\boldsymbol{v}}^{(\mathbf{1})}$ sono le approssimazioni calcolate per λ_1 e $\boldsymbol{v}^{(\mathbf{1})}$.